DIAL - Enhanced Interoperability

A Distributed Catalog and Data Services System for Remote

Software tools based on WWW technology have greatly enhanced access to remote sensing and global environmental data. These tools provide users capabilities to quickly preview data for important disaster related events such as forest fire, volcanic events, floods, storms, hurricanes, using a web browser. Users will also be able to see time series events such as ozone level changes, vegetation index changes, land cover and land use changes from a variety of instruments in multiple satellite missions launched by many space agencies in the world. On the following pages we are presenting one of NASA's efforts to provide and enhance data access to satellite and other types of remote sensing data based on WWW technologies through Data and Information Access Link (DIAL).

By Ramachandran Suresh, Liping Di, Lakshmi Kumar, Kenneth McDonald

Data and Information Access Link (DIAL) is a web based package of software tools for Remote sensing and geo-spatial data applications. Using DIAL, scientists can set up a web based data server, organize data, build metadata catalog, and distribute data. Users with the help of a web browser can access a DIAL site, search and query data, browse metadata, visualize data, perform animation, subset and subsample data, and download data on-line in multiple formats. DIAL has implemented EOS Data Gateway (EDG) catalog protocol and the Committee on Earth Observation Satellites (CEOS) Catalog Interoperability Protocol (CIP) through which projects can set up a distributed data system. The use of this system enhances data interoperability.

Low-Cost Web-Based Solution

In addition to providing easy access to remote sensing satellite data to users, DIAL also provides a mechanism for data producers to set up a system using DIAL to distribute, share and provide access to their own collection of data. Before the advent of WWW technology, the process of setting up a data system or center to distribute and exchange scientific data was a difficult and expensive process. DIAL has made it possible to set up low cost data distribution systems for satellite remote sensing and field campaign data. DIAL enables scientists and projects to set up metadata catalogs, provide browsing, visualization and on-line downloading of data. DIAL makes it easy for scientists to collaborate and exchange scientific data on-line. NASA funded the DIAL project to provide a low-cost web-based solution for managing and distributing Earth science remote sensing and geo-spatial data. The new technologies have made it possible for small data producers and projects to set up such systems on their desktop and provide data and catalog services. The DIAL system described here can also support the setting up of distributed data systems that can inter-operate and exchange scientific data through the implementation of certain standard protocols.

Before the advent of WWW technology, the process of setting up a data system or center to distribute and exchange scientific data was a difficult and expensive process.

DIAL Description

DIAL is a web-based software system for distributing scientific data through the Internet (Di et al., 1997, 98, McDonald et al.,1998). DIAL is designed to be a client-server based data and information system that, while powerful, is also compact, easy to setup and use, and has minimal computer power and maintenance requirements. It permits data producers to set up a server rapidly on their desktop computers, making data available via the Internet. Any users with a regular Web browser can interactively search, browse, subset, subsample, reformat, and download data.

Functionality

The DIAL system provides two major types of user services: the catalog service and the data service. The catalog service allows data users to find individual data objects (granules) in an archive by entering some search criteria. Once a user finds the needed individual data objects in the archive, the data service will allow the user to manipulate and download the objects. The DIAL software is available for users free of charge at http://dial.gsfc.nasa.gov. A user can visit the DIAL site, run a demo, and visualize the features described in this paper.

Catalog and Data Services

Catalog services in DIAL include:

- Spatial, temporal, and parameter-based search
- Catalog search at inventory and directory levels
- Automatic creation of catalog based on metadata
- Supports of ODBC/JDBC compatible databases for storing metadata/catalog
- User-friendly Java search interface with query preview
- HTML search interface for slow network connection

# of Geo-spatial Data

Sensing Data

- Distributed Data Service (Search catalog can support remote datasets)

Data services in DIAL include:

- On-line access to data and metadata
- Single- and multi-granule subsetting and subsampling based on array coordinates or record numbers
- Single- and multi-granule subsetting and subsampling based on geographic/map coordinates and physical parameters
- Browse and dynamic/interactive visualization of data
- Single- and multi-granule animation of time series or high dimensional data
- Interactive color composites of multi-spectral data
- Overlay of coastal and political boundaries
- On-line downloading of data in multiple formats
- X-Y plotting for non-image data

Animation

Users can animate time series data. Figure 1 shows an example of ozone data animation. This example has 30 days of data in the imagery form. A user can select the animation option and palette. A user can stop at a particular frame, play the animation, go to a previous image or can see either the first or the last image. A user can also set display duration. Users can also overlay political and coastline boundaries on top of the image. Figure 1 shows the coastline in

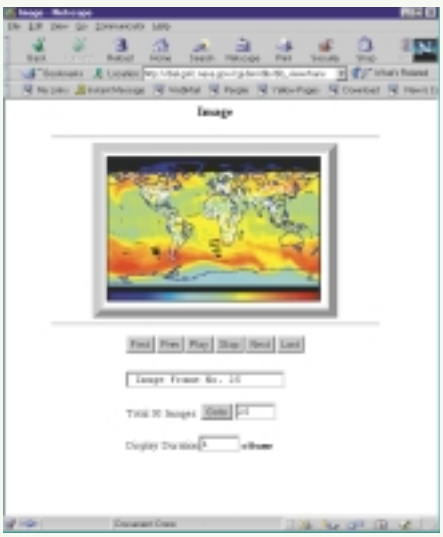

Figure 1: an example of ozone time series data animation.

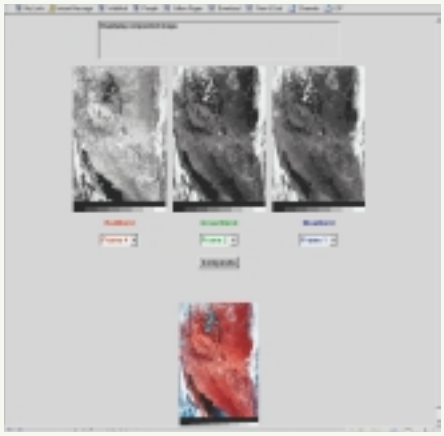

Figure 2: an example of color composite image from MODIS data with three-band color composite function of DIAL.

white and the political boundaries in blue. The animation function of DIAL is very useful to identify temporal changes in data.

The Capability of Color Composite

A user can view multispectral data from different bands with three-band color composite function of DIAL. Figure 2 shows an example of color composite image from MODIS data. When a user selects the pseudo color composite option, a color composite client with three band-image windows and one composite image window will appear on the user's web. Then user can select any one of the bands from the window below each image. After the selection if the user clicks the composite button, a color composite of the image from three bands representing RGB appears on the screen.

Plotting

DIAL allows users to visualize tabular data stored in HDF or HDF-EOS files. Users can perform multivariate plotting of the data by selecting fields from the table for X-Y axis. A user can plot all the data points or can plot only a sub set of the data. Legend can be created and graph can be labeled. Figure 3 shows the x-y plot of correction factors.

Sub-sampling and Sub-setting

Figure 4 and 5 shows an example of sub-sampling and sub-setting of data. This feature is particularly useful because many of the satellite imagery data sets such as those from MODIS, Landsat7, CERES, ASTER, and MISR, are very large (more than 200 MB). The user may select every third or fourth pixel of the image (sub-sampling) and/or select the start and end points of the pixel dimensions of the image (sub-setting) for display and downloading.

On-line Downloading of Data in Multiple Formats

A user can download data in HDF, binary or ASCII formats regardless what is the data format in the server. User can also view the data on screen. A user can download an entire data file or a subset of the data file and do further analysis by using specialized data analysis packages. DIAL currently works with HDF, HDF-EOS, HDF5 and netCDF (IRIX, OSF1) formatted data sets. Other formats are supported through the use of translation programs. Commonly-used GIS formats such as Arc/Info Exchange

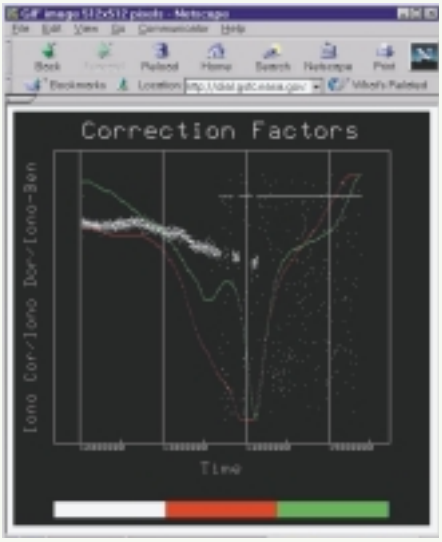

Figure 3: an example of the x-y plot of correction factors.

Format, GeoTIFF, ERDAS LAN, and ArcView Shape are supported through the use of translation programs. In addition, DIAL's architecture allows extension for accessing data in other formats.

Storing the Searchable Catalog

DIAL can work with helper applications such as WebWinds (http://webwinds.jpl.nasa.gov). A user can configure his/her browser to use WebWinds as a plug-in application for HDF file types. The WebWinds application gets launched when a datafile is downloaded

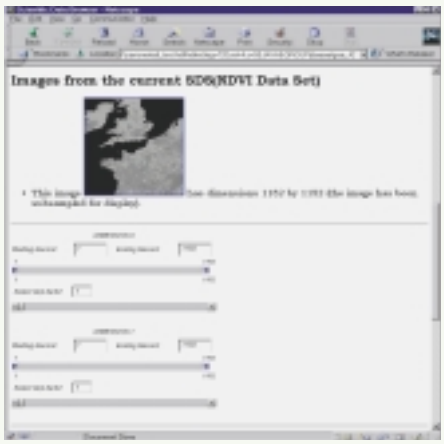

Figure 4): an example of sub-sampling of data

from a DIAL site in HDF format, allowing the user to do further data analysis.
DIAL provides two options to store the searchable catalog:

1) a binary table on which DIAL's own search engine will search directly; and
2) ODBC-compatible databases on which DIAL will search through its generic JDBC interface. The binary table option provides the same search capabilities as the ODBC-compatible databases, but the search speed will be slowed down when the table is very large.

Distributed DIAL System

DIAL can be used in two different ways: either as a standalone data server or as a distributed system. DIAL can work as a distributed system connecting several DIAL sites through the EOS Data Gateway (EDG), (Suresh et al, 1999). DIAL supports EOSDIS "Version 0" protocol, extending its catalog interoperability to EOSDIS user community. EDG is a distributed information management system that can search multiple distributed data catalogs through its query interface. EDG provides a consistent view of more than 900 data products held at each of the EOSDIS DAACs and several international data centers. Users without specific knowledge of the data can search science data holdings, retrieve detailed data inventories and high-level descriptions of data sets, view sample browse imagery, and place orders for data products. Both free-text simple search and advanced data attribute search interfaces are available. Data descriptions of NASA Global Change Master Directory (GCMD) are integrated into the Directory search function, improving the consistency and efficiency of

Figure 5: an example of sub-setting of data

providing useful extension of the information in the system.

EDG

The goal of EDG is to facilitate Earth science research through improved access to existing data and to serve as

NASA funded the DIAL project to provide a low-cost web-based solution for managing and distributing Earth science remote sensing and geo-spatial data.

the search and order tool for data products produced by EOS. EDG uses Version 0 protocols and a data dictionary for providing catalog interoperability. Recently, EDG has been assigned a new role as the operational search and order tool for Landsat-7, Terra and other data sources. This service is available at http://eos.nasa.gov/imswelcome
DIAL will be supporting Catalogue Interoperability Protocol (CIP) in the near future (March 2001) as it is the de-facto international standard for catalog interoperability of geo-spatial data. Supporting CIP protocol will enhance the data interoperability to a wider community of remote sensing data. Figure 7 shows the supported protocols in DIAL, including HTTP, Vo and CIP.

Architecture

The DIAL architecture is modular, extendible, and is based on standards. Figure 8 shows the architecture. The DIAL system consists of three components: the end user interface, the DIAL server system, and the data management tools. The end user interface on the client side includes web browser with server-generated Java or HTML interfaces for querying the system to locate datasets of interest and for interactively manipulating the datasets. Requests from a user are relayed to DIAL server system through http protocols. DIAL end users can easily integrate other application software as helper applications. The DIAL server system is CGI-based programs which respond to users' requests and send the reply to the end user interface through an HTTP server. The DIAL server system works with the data in HDF, HDF5, HDF-EOS, and netCDF formats. The metadata is stored in a binary table or a database either in Object Description Language (ODL) or Parameter Value Language (PVL) format. The interoperability layer

The DIAL software is available for users free of charge at http://dial.gsfc.nasa.gov.

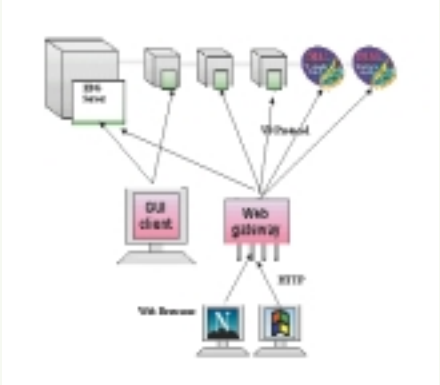

Figure 6: Earth Science Data Gateway (EDG)

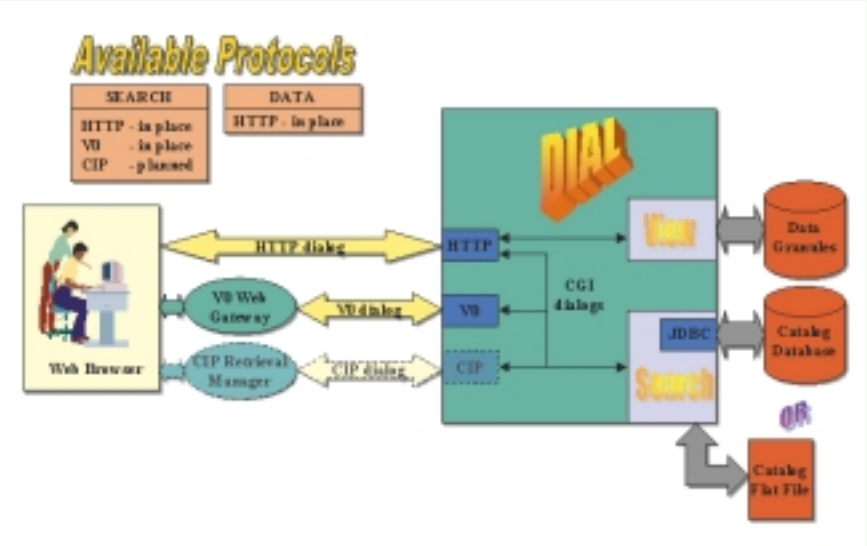

Figure 7: The supported protocols in DIAL, including HTTP, V0 and CIP.

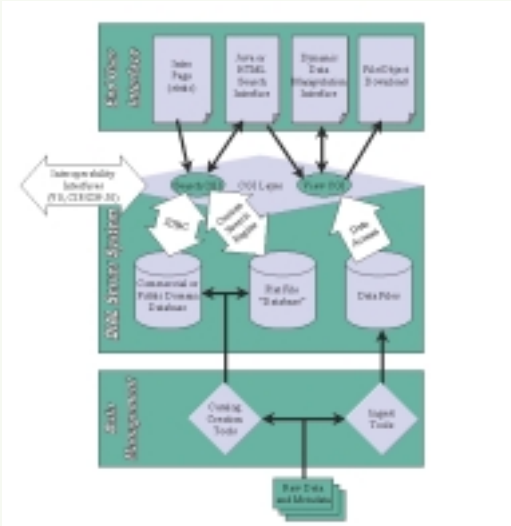

Figure 8: DIAL architecture

in the server provides interoperations with other data systems through standards. The data management tools help data providers to prepare the data and metadata for distribution through the DIAL system.

DIAL uses WWW standards such as http, html and Java applets in implementation. It can work with any metadata standards, such as FGDC, since its metadata catalog system is totally configurable and a mapping mechanism between different metadata schema is provided. The DIAL software package is compact (less than 7 Megabytes) and does not need programming knowledge to use. Since it is completely developed from public domain software and protocols, it is not dependent on any COTS package. It is available on multiple platforms such as SGI, SUN, DEC, and PCs (Windows 95/98/NT). The source code is written in C and Java.

A few thousand users have received DIAL software on CD-ROMs and many more have downloaded the software from the web site.

Users

A few thousand users have received DIAL software on CD-ROMs and many more have downloaded the software from the web site. Some of the well known DIAL installations include NASA EOS Principal Investigators, NASA Earth Science Information Partners (ESIP), National Oceanic and Atmospheric Administration's Pacific Marine Environmental Laboratory, Jet Propulsion Laboratory, and international partners (Japan, United Kingdom, China, Australia, European Space Agency, IGBP START South East Asian and East Asia Regional Centers).

Conclusions

Web based data systems can take advantage of standard protocols to create a distributed system that enhances scientific data exchange and interoperability. DIAL system with EDG and CIP provides a distributed system for users and can bring the abundance of various remote sensing data right to users' desktops.

Ramachandran Suresh (suresh@rattler.gsfc.nasa.gov) is the Manager of the Data Eng. Section at Raytheon ITSS working on a NASA contract. His research interests include WWW technology, information standards, remote sensing and GIS applications.
Dr. Liping Di (lpd@rattler.gsfc.nasa.gov) is a chief scientist with Raytheon ITSS. His research interests include interoperability of data and system, geographic information standards, and remote sensing applications.
Lakshmikantha Kumar (lakshmi@rattler.gsfc.nasa.gov)is a Senior Software Engineer at Raytheon ITSS working for the DIAL project. She has designed and developed software for many other satellite missions such as SeaWiFS and EOS Terra.
Kenneth McDonald (kenneth.mcdonald@gsfc.nasa.gov) works for NASA ESDIS Project at Goddard Space Flight Center as Manager for external development and interfaces office. He was involved in the development of EDG and DIAL and the chair of CEOS WGISS Access Subgroup.

Acknowledgments
This research is funded by NASA Technology prototype under NASA contract NO: 50-007

References
CEOS/WGISS/PTT/CIP-B, June 1998 Issue 2.4, Catalog Interoperability Protocol (CIP) Specification - Release B, Committee on Earth Observation Satellites
Specification, ANSI/NISO Z39.50-1995, Official Text for Z39.50-1995, July 1995 Information Retrieval (Z39.50): Application Service Definition and Protocol
Suresh, R., Di. L., Kumar.L., Ilg.D.,McDonald, K., 1999. A Web-based Access to Distributed Heterogeneous Geospatial Data Archives using NASA's EOSDIS V0 Protocol. In Chen, J., Zhou, Q., Li Zhilin., Jiang Jie (eds) Dynamic and Multidimensional GIS, Vol 32, Part 4W12, Beijing, China, P 371.
Di, L, Suresh,R., Doan, K., Ilg. D., McDonald, K.,1998. DIAL-an Interoperable Web-based Scientific Data Distribution System. In M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman (eds), Interoperating Geographic Information Systems, Section 4. System Experiences. Kluwer Academic Publishers, Norwell, MA.
Di, L, Suresh, R., Doan, K., Ilg, D., McDonald, K.,1997. A Web-based Scientific Data Server for Accessing and Distributing Earth Science Data. In Proceedings of Interop'97 PP 250-261
McDonald, K., Suresh, R., Di. L., Kumar. L 1998, Data and Information Access Link (DIAL), Fall meeting American Geophysical Union